

# Review of recent results on optimal orthonormal subband coders

P. P. Vaidyanathan, California Institute of Technology, Pasadena, CA 91125

email: ppvnath@sys.caltech.edu

**DISTRIBUTION STATEMENT 2**

Approved for public release  
Distribution Unlimited

## ABSTRACT

The conditions for optimality of uniform orthonormal subband coders are reviewed. A number of properties of optimal filter banks are then summarized. The case of nonuniform orthonormal filter banks is also considered, and it is shown that the well known connection between optimal coding gain, energy compaction, and principal component property does not extend to the nonuniform case.

**Keywords:** Optimal subband coders, orthonormal filter banks, energy compaction.

## 1. INTRODUCTION

The optimization of orthonormal filter banks<sup>1,5</sup> for fixed input statistics has been of some interest in recent years<sup>6-18</sup>. Similar to classical transform coders, the optimality criterion is usually the coding gain (so that the mean square reconstruction error due to subband quantization is minimized). Some of the advances in recent literature have addressed the theory of **uniform** filter banks (i.e., identical decimation ratio in all channels) from the contexts of energy compaction<sup>8,9</sup>, and principal component analysis<sup>10,14</sup>. Numerical optimization of the coding gain has also been addressed<sup>7,11,12,13</sup>. A set of necessary and sufficient conditions for optimality was developed in<sup>16</sup>, and it was shown that these conditions can be satisfied by the design of a sequence of compaction filters. The uniform orthonormal filter bank optimized for coding gain is also the principal component filter bank, with each filter acting as an energy compaction filter for an appropriate power spectrum related to the input. In this tutorial we first review these results, and then consider the case of nonuniform orthonormal filter banks. For this case we will see that there is no simple connection between coding gain, energy compaction, and the principal component property.

### Preliminaries

The subband coder shown in Fig. 1(a) is said to be maximally decimated if the decimation ratios  $n_k$  satisfy  $\sum_k 1/n_k = 1$ . It is a uniform subband coder if the  $n_k = M$  for all  $k$ . All discussions except those in Sec. 4 are restricted to this case. In this case we have the polyphase representation<sup>2</sup> shown in Fig. 1(b). The filter bank is said to be a biorthogonal or perfect reconstruction filter bank if  $\mathbf{R}(z)\mathbf{E}(z) = \mathbf{I}$ . It is said to be orthonormal or paraunitary if  $\mathbf{E}(e^{j\omega})$  is unitary for all  $\omega$ . In the orthonormal case the perfect reconstruction condition is  $F_i(e^{j\omega}) = H_i^*(e^{j\omega})$ . In terms of the filters we can express orthonormality as<sup>2</sup>  $H_k(e^{j\omega})H_m^*(e^{j\omega})|_{\downarrow M} = \delta(k-m)$  where the notation  $F(e^{j\omega})|_{\downarrow M}$  denotes the Fourier transform of  $f(Mn)$ . The orthonormality condition implies in particular that each filter  $H_i(e^{j\omega})$  satisfies the **Nyquist(M) constraint**  $|H_i(e^{j\omega})|^2|_{\downarrow M} = 1$ . Traditional contiguous-stackings of brickwall filters (Fig. 2) serve as examples of orthonormal filter banks. In these two examples each filter is an aliasfree(M) filter. An **aliasfree(M)** or **antialias(M)** filter is defined to be one whose output can be decimated without aliasing, that is, the shifted versions  $H_i(e^{j(\omega-2\pi k/M)})$  do not overlap for distinct  $k$  in  $0 \leq k \leq M-1$ . Equivalently we say that the filters have aliasfree(M) supports. Such a support could have

multiple number of passbands (e.g., see Fig. 7 later.)

**Statistical Model.** The input  $x(n)$  is assumed to be zero-mean wide sense stationary (WSS) with power spectral density (psd)  $S_{xx}(e^{j\omega})$ . The subband signals  $v_i(n)$  (see Fig. 1(a)) are therefore (zero-mean and) jointly WSS, and the variances of  $x_i(n)$  and  $v_i(n)$  are the same, that is,  $\sigma_{x_i}^2 = \sigma_{v_i}^2$ . We model the quantizers with additive noise sources  $q_i(n)$  (Fig. 1(a)), and assume these noise sources to be jointly WSS with zero mean and variances of the form<sup>2,19</sup>  $\sigma_{q_i}^2 = c2^{-2b_i}\sigma_{x_i}^2 = c2^{-2b_i}\sigma_{v_i}^2$  where  $b_i$  is the number of bits assigned to the  $i$ th subband quantizer. The constant  $c$  (which depends on the nature of the pdf of the quantizer input<sup>19</sup> is assumed to be the same for all subbands. This model does not require that each  $q_i(n)$  be white or that any two noise sources be uncorrelated.

**Coding Gain.** The quantity  $b = \sum_{i=0}^{M-1} b_i/M$ , which is the **average bit rate**, is assumed to be fixed. The coding gain of a subband coder is defined by comparing the average mean square value  $\mathcal{E}_{SBC}$  of the reconstruction error  $x(n) - \hat{x}(n)$  with the m.s. value  $\mathcal{E}_{direct}$  of the direct quantization error (roundoff quantizer) with the same bit rate  $b$ . Using the above noise model, the coding gain  $G_{SBC}(M)$  of the uniform orthonormal subband coder of Fig. 1(a) can be derived<sup>2</sup>:

$$G_{SBC}(M) \triangleq \frac{\mathcal{E}_{direct}}{\mathcal{E}_{SBC}} = \frac{\sum_{i=0}^{M-1} \sigma_{x_i}^2 / M}{\left(\prod_{i=0}^{M-1} \sigma_{x_i}^2\right)^{1/M}} = \frac{\sigma_x^2}{\left(\prod_{i=0}^{M-1} \sigma_{x_i}^2\right)^{1/M}} \quad (1)$$

This expression assumes *optimal bit allocation* as described in<sup>2,19</sup>. Here we have used the result  $\sum_{i=0}^{M-1} \sigma_{x_i}^2 = M\sigma_x^2$ , valid for uniform orthonormal filter banks. The preceding coding gain is the ratio of the arithmetic and geometric means, **AM/GM ratio**, of the subband variances  $\sigma_{x_i}^2$ . For fixed input psd  $S_{xx}(e^{j\omega})$  the variances  $\sigma_{x_i}^2$  depend only on the analysis filters  $H_i(e^{j\omega})$ . We say that the **subband coder is optimal** (for the fixed  $M$  and  $S_{xx}(e^{j\omega})$ ), if these filters are such that the coding gain is maximized. From (1) we see that coding gain optimization is equivalent to minimizing the **product of subband variances**. If the psd  $S_{xx}(e^{j\omega})$  is monotone decreasing then the contiguous stacking (Fig. 2) yields an optimal orthonormal filter bank. For arbitrary psd, each filter in the optimal system can have many passbands<sup>16</sup>.

## 2. OPTIMALITY OF UNIFORM ORTHONORMAL FILTER BANKS

We now present the conditions for the optimality of an uniform orthonormal filter bank.

### Total Decorrelation Of Decimated Subbands

In orthogonal transform coding theory (where  $\mathbf{E}(z)$  in Fig. 1(b) is a constant unitary matrix), decorrelation of the decimated subband random variables  $v_i(n)$  is necessary and sufficient for optimality. That is, the coding gain is maximum if and only if  $E[v_i(n)v_k^*(n)] = 0$  for  $i \neq k$ , and for all  $n$ . For orthonormal subband coders, a stronger condition is necessary: namely, the *decimated subband random processes* should be uncorrelated, that is,

$$E[v_i(n)v_k^*(m)] = 0 \quad (2)$$

for  $i \neq k$ , and for all  $n, m$ . This condition will also be referred to as **total decorrelation** of subbands. Equivalently, the power spectrum matrix of the vector process  $\mathbf{v}(n) \triangleq [v_0(n) \ v_1(n) \ \dots \ v_{M-1}(n)]^T$  must be diagonal. That is, letting  $S_i(e^{j\omega})$  be the psd of  $v_i(n)$ , we should have  $\mathbf{S}_{\mathbf{v}\mathbf{v}}(e^{j\omega}) = \text{diag}[S_0(e^{j\omega}) \ S_1(e^{j\omega}) \ \dots \ S_{M-1}(e^{j\omega})]$ . *Even this stronger condition is not sufficient* for optimality. For example the traditional brickwall subband coder in Fig. 2 decorrelates the

subband processes trivially for any input process. The preceding necessary condition follows from the fact that if a pair of decimated subband processes, say  $v_0(\cdot)$  and  $v_1(\cdot)$ , are not uncorrelated, then we can insert a delay  $z^{-k}$  and a unitary matrix  $\Theta$  to transform the pair  $v_0(n), v_1(n-k)$  into an uncorrelated pair  $w_0(n)$  and  $w_1(n)$  (Fig. 3). The modified filter bank continues to be orthonormal. It can be shown that this modification is such that  $\sigma_{w_0}^2 \sigma_{w_1}^2 < \sigma_{v_0}^2 \sigma_{v_1}^2$  which shows that the coding gain has been increased. Thus, if total decorrelation is not true, the coding gain can be increased. Note that total decorrelation of the undecimated subbands  $x_k(n)$  is not necessary.

Using the preceding necessary condition, we can show that the optimal uniform orthonormal subband coder has  $G_{SBC}(M) = 1$  if and only if the (zero-mean WSS) input process  $x(n)$  has power spectral density of the form  $S_{xx}(e^{j\omega}) = S(e^{j\omega M})$ . In this case all the decimated subband signals  $v_i(n)$  have identical psd  $S(e^{j\omega})$ . The condition  $S_{xx}(e^{j\omega}) = S(e^{j\omega M})$  is equivalent to the condition that the autocorrelation  $R_{xx}(k)$  of  $x(n)$  is nonzero only when  $k$  is a multiple of  $M$ . Thus, there exist **nonwhite** inputs for which orthonormal SBC might yield no gain, for some fixed  $M$ . However, unless the input is white, this will not happen for all values of  $M$ .

### Majorization of Decimated Subbands

We say that the set of decimated subband signals  $v_k(n)$  has the **majorization property** if their power spectra  $\{S_k(e^{j\omega})\}$  satisfy

$$S_0(e^{j\omega}) \geq S_1(e^{j\omega}) \geq \dots \geq S_{M-1}(e^{j\omega}), \quad \text{for all } \omega. \quad (3)$$

That is, the  $k$ th subband psd dominates the  $(k+1)$ th psd for all  $\omega$ . This is demonstrated in Fig. 4.

**Theorem 1. Majorization is necessary.** For fixed input psd  $S_{xx}(e^{j\omega})$  and fixed  $M$ , suppose a subband coder is optimal (in the coding gain sense) among the class of all  $M$ -band uniform orthonormal subband coders. Then the decimated subband signals  $v_k(n)$  have the majorization property.  $\diamond$

**Proof.** Assume majorization is not satisfied, e.g., let  $S_0(e^{j\omega}) \geq S_1(e^{j\omega})$  be not valid for all  $\omega$ , even though  $\sigma_{v_0}^2 \geq \sigma_{v_1}^2$ . Cascade a new paraunitary matrix  $\mathbf{T}(e^{j\omega})$  to the right of the given filter bank as shown in Fig. 5, where

$$\mathbf{T}(e^{j\omega}) = \begin{cases} \mathbf{I}_M & \text{if } S_0(e^{j\omega}) \geq S_1(e^{j\omega}) \\ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & \mathbf{I}_{M-2} \end{bmatrix} & \text{if } S_0(e^{j\omega}) < S_1(e^{j\omega}) \end{cases} \quad (4)$$

The new pair of power spectra  $S'_0(e^{j\omega}), S'_1(e^{j\omega})$  will then satisfy the property  $S'_0(e^{j\omega}) \geq S'_1(e^{j\omega})$  for all  $\omega$ . Moreover for each  $\omega$ ,  $S'_0(e^{j\omega}) \geq S_0(e^{j\omega})$  whereas  $S'_1(e^{j\omega}) \leq S_1(e^{j\omega})$ . Thus the variances of the new signals  $w_0(n)$  and  $w_1(n)$  are such that  $\sigma_{w_0}^2 \geq \sigma_{v_0}^2$  and  $\sigma_{w_1}^2 \leq \sigma_{v_1}^2$ . But  $\sigma_{w_0}^2 + \sigma_{w_1}^2 = \sigma_{v_0}^2 + \sigma_{v_1}^2$  (since  $\mathbf{T}(e^{j\omega})$  is paraunitary). Using this we can verify  $\sigma_{w_0}^2 \sigma_{w_1}^2 < \sigma_{v_0}^2 \sigma_{v_1}^2$ . So the coding gain can be increased whenever majorization is not satisfied.  $\nabla \nabla \nabla$

### Necessary and Sufficient Condition For Optimality

Though majorization and decorrelation are necessary for optimality, neither of them is individually sufficient. For example, the brickwall subband coder with contiguous stacking (e.g., Fig. 2) satisfies the total decorrelation property for any input psd. On the other hand the orthonormal system with  $H_k(z) = z^{-k}$  satisfies majorization for any input though it yields no coding gain. However we have<sup>16</sup>:

**Theorem 2. A necessary and sufficient condition for optimality.** For fixed input psd  $S_{xx}(e^{j\omega})$ , the filter bank has maximum coding gain (among all  $M$  band uniform orthonormal subband coders) if and only if the decimated subband signals  $v_k(n)$  *simultaneously* satisfy total decorrelation and majorization. Furthermore, when these conditions are satisfied, the set of power spectra  $\{S_k(e^{j\omega})\}$  of the decimated subband signals is unique.  $\diamond$

**Proof.** From Fig. 1(b) we have  $\mathbf{S}_{vv}(e^{j\omega}) = \mathbf{E}(e^{j\omega})\mathbf{S}_B(e^{j\omega})\mathbf{E}^\dagger(e^{j\omega})$  where  $\mathbf{S}_B(e^{j\omega})$  and  $\mathbf{S}_{vv}(e^{j\omega})$  are power spectra of the vectors  $\mathbf{x}_B(n)$  and  $\mathbf{v}(n)$  indicated in the figure. If  $\mathbf{E}(e^{j\omega})$  performs total decorrelation,  $\mathbf{S}_{vv}(e^{j\omega})$  is diagonal:

$$\mathbf{S}_{vv}(e^{j\omega}) = \mathbf{E}(e^{j\omega})\mathbf{S}_B(e^{j\omega})\mathbf{E}^\dagger(e^{j\omega}) = \text{diag} \{S_0(e^{j\omega}) S_1(e^{j\omega}) \dots S_{M-1}(e^{j\omega})\} \quad (5)$$

Since  $\mathbf{E}(e^{j\omega})$  is unitary, this implies that for each fixed  $\omega$ , the subband power spectra  $S_k(e^{j\omega})$  are **eigenvalues** of  $\mathbf{S}_B(e^{j\omega})$ . Suppose the majorization property also holds. This means that for each  $\omega$  these eigenvalues are ordered in a decreasing fashion. Since the set of eigenvalues is unique, each diagonal element in (5) is uniquely determined for each  $\omega$ . Thus the set of power spectra  $\{S_k(e^{j\omega})\}$  which has the majorization property is unique as claimed in the theorem. Since majorization and total decorrelation are necessary for optimality, and since there is only one set of majorized decorrelated subband power spectra, it follows that majorization together with decorrelation leads to optimality.  $\nabla \nabla \nabla$

Since the diagonalizing eigenvector matrix may not be unique, the analysis filters of the optimal system may not be unique.

### 3. COMPACTION FILTERS AND CODING GAIN MAXIMIZATION

Fig. 6 shows a filter  $H(e^{j\omega})$  with a zero-mean WSS input  $x(n)$  having psd  $S_{xx}(e^{j\omega})$ . This can be regarded as an  $M$ -fold decimation filter, that is, one branch of an  $M$ -channel analysis bank. Consider the problem of designing  $H(e^{j\omega})$  such that the output variance  $\sigma_y^2$  is maximized subject to the constraint that  $|H(e^{j\omega})|^2$  be Nyquist( $M$ ), that is,

$$|H(e^{j\omega})|^2 \Big|_{\downarrow M} = 1, \quad \text{i.e.,} \quad \sum_{k=0}^{M-1} |H(e^{j(\omega - 2\pi k/M)})|^2 = M, \quad \text{for all } \omega \quad (\text{Nyquist constraint}). \quad (6)$$

The solution  $H(e^{j\omega})$  is an **optimum compaction(M) filter**. The ratio  $\sigma_y^2/\sigma_x^2$  is called the **compaction gain**. The Nyquist constraint is imposed because  $|H_k(e^{j\omega})|^2$  is Nyquist( $M$ ) for every analysis filter  $H_k(e^{j\omega})$  in any uniform orthonormal filter bank. This constraint implies the **unit-energy property**  $\int_0^{2\pi} |H(e^{j\omega})|^2 d\omega/2\pi = 1$  as well as the **boundedness property**  $|H(e^{j\omega})|^2 \leq M$ . The role of the energy compaction concept in subband coding theory has been observed by a number of authors<sup>8-14</sup>. The following is a refined version for arbitrary  $M$ , of Unser's construction of compaction filters<sup>8</sup>: (a) For each frequency  $\omega_0$  in  $0 \leq \omega < 2\pi/M$  define the  $M$  alias frequencies  $\omega_k = \omega_0 + 2\pi k/M$ , where  $0 \leq k \leq M-1$ . (b) Compare the values of  $S_{xx}(e^{j\omega})$  at these  $M$  alias frequencies  $\{\omega_k\}$ . Let  $L$  be the smallest integer such that  $S_{xx}(e^{j\omega_L})$  is a maximum in this set. Then assign

$$H(e^{j(\omega_0 + (2\pi k/M))}) = \begin{cases} \sqrt{M} & \text{when } k = L \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

Repeating this for each  $\omega_0$  in the region  $0 \leq \omega < 2\pi/M$ , the filter  $H(e^{j\omega})$  is completely defined for all  $\omega$  in  $0 \leq \omega < 2\pi$ . This filter maximizes the output variance  $\sigma_y^2$  under the Nyquist( $M$ ) constraint (6).

**Remarks.** If  $H(e^{j\omega})$  is an optimal compaction( $M$ ) filter for an input psd  $S_{xx}(e^{j\omega})$  then it will be a valid optimal solution for the modified psd  $f[S_{xx}(e^{j\omega})]$  where  $f[\cdot] \geq 0$  is any nondecreasing function. If a psd is non increasing in  $[0, 2\pi)$ , then the optimum compaction filter is lowpass. While the optimal compaction filter is not unique, the construction described above has the following special properties: (a)  $H(e^{j\omega})$  is an ideal **two-level filter** with passband response  $= \sqrt{M}$  and stopband response  $= 0$ . (b)  $H(e^{j\omega})$  is an **antialias( $M$ ) filter**. This follows because if  $H(e^{j\omega}) \neq 0$  for some  $\omega$ , then  $H(e^{j(\omega-2\pi k/M)}) = 0$  for  $1 \leq k \leq M-1$ . (c) The **total width of all passbands** is  $2\pi/M$ .

To describe the construction of optimum (uniform orthonormal) filter banks, consider the example of input psd shown in Fig. 7(a), and let  $M = 4$ . The first step is to choose one filter,  $H_0(e^{j\omega})$ , to be an optimal energy compaction filter for the input psd  $S_{xx}(e^{j\omega})$ . This filter is shown in Fig. 7(b). Let the passband support of  $H_0(e^{j\omega})$  be denoted  $\mathcal{S}_0$ . Suppose we define a new psd

$$S_{xx}^{(1)}(e^{j\omega}) = \begin{cases} 0 & \text{in } \mathcal{S}_0 \\ S_{xx}(e^{j\omega}) & \text{otherwise} \end{cases} \quad (8)$$

as shown in Fig. 7(c). Thus  $S_{xx}^{(1)}(e^{j\omega})$  is obtained by **peeling off** the portion of  $S_{xx}(e^{j\omega})$  falling in the passband of  $H_0(e^{j\omega})$ . We design the next analysis filter  $H_1(e^{j\omega})$  to be the optimal compaction filter for this partial psd  $S_{xx}^{(1)}(e^{j\omega})$ . This is shown in Fig. 7(d). Define the next partial psd  $S_{xx}^{(2)}(e^{j\omega})$  by peeling off the portions of  $S_{xx}(e^{j\omega})$  in the passbands of  $H_0(e^{j\omega})$  and  $H_1(e^{j\omega})$ , and continue in this manner. Thus all the analysis filters can be identified (part (e) in the figure). The proof that this construction results in an optimal orthonormal system can be found in<sup>17</sup>. Most readers will recognize that the preceding construction yields a **principal component filter bank**<sup>10</sup>. This connection between coding gain, principal component property, and compaction gain does not carry over to the case of nonuniform filter banks (Sec. 4).

In general the optimum compaction filters are ideal infinite order filters. In practice, if we approximate these with FIR filters rather than approximating Fig. 2 as done traditionally, the coding gain can be significantly improved. It can be shown that the maximum compaction gain is unity for an input psd  $S_{xx}(e^{j\omega})$  (for a fixed value of  $M$ ) if and only if it has the form  $S_{xx}(e^{j\omega}) = S(e^{j\omega M})$ . In fact, the following three statements are equivalent<sup>17</sup>: (a)  $S_{xx}(e^{j\omega})$  has the form  $S_{xx}(e^{j\omega}) = S(e^{j\omega M})$ . (b) Maximum compaction gain  $G_{max}(M) = 1$ . (c) Any  $M$ -band orthonormal SBC yields coding gain  $= 1$ .

#### 4. NONUNIFORM FILTER BANKS

The nonuniform filter bank of Fig. 1(a) is orthonormal<sup>20</sup> if the analysis filters satisfy  $H_k(e^{j\omega})H_m^*(e^{j\omega})|_{g_{km}} = \delta(k-m)$  where  $g_{km} = gcd\{n_k, n_m\}$ . The coding gain of the nonuniform  $M$ -band orthonormal filter bank has been derived in<sup>20</sup> and is given by

$$G_{SBC}(M) = \frac{\sigma_x^2}{\prod_{i=0}^{M-1} (\sigma_{x_i}^2)^{1/n_i}} \quad (9)$$

Consider the special case of a two-level tree structured filter bank (Fig. 8), which is equivalent to a 3-channel nonuniform FB. The coding gains  $G_1$  and  $G_2$  of the two-channel filter banks at the levels 1 and 2 are given by  $G_1 = \sigma_x^2 / \sqrt{\sigma_{y_1}^2 \sigma_{x_2}^2}$  and  $G_2 = \sigma_{y_1}^2 / \sqrt{\sigma_{x_0}^2 \sigma_{x_1}^2}$ . Thus the coding gain of the three channel filter bank is

$$G_{SBC}(3) = \frac{\sigma_x^2}{(\sigma_{x_2}^2)^{1/2} (\sigma_{x_1}^2)^{1/4} (\sigma_{x_0}^2)^{1/4}} = G_1 G_2^{1/2}$$

More generally, consider the  $M$  channel dyadic tree structure with  $M - 1$  levels (Fig. 9(a)). Assuming each two channel filter bank is orthonormal, we have an  $M$ -band nonuniform orthonormal filter bank, and the coding gain is<sup>18</sup>

$$G_{SBC}(M) = G_1 G_2^{1/2} G_3^{1/4} \dots \quad (10)$$

where  $G_m$  is the coding gain at level  $m$ . Thus the extra benefit offered by the  $m$ th split decays exponentially with  $m$ . Another consequence of (10) is that, if the tree structure maximizes the coding gain for a given input then the *right-flushed subtrees indicated in Fig. 10 are optimal for their respective inputs*. However, since these inputs depend on the filters preceding them, this observation is not directly useful to identify the optimal filter bank.

### Coding Gain and Compaction Gain

Consider the input power spectrum shown in Fig. 11(a). Assuming that the process is Gaussian, the rate-distortion theoretic upper bound on the coding gain, given by<sup>19</sup>  $G_{up} = \sigma_x^2 / \exp \int_0^{2\pi} \text{Ln } S_{xx}(e^{j\omega}) d\omega / 2\pi$ , has the value

$$G_{up} = \frac{2 + c + d}{4(cd)^{1/4}} \quad (11)$$

Suppose we use the two-level orthonormal tree structured filter bank shown in Fig. 8, with filter responses as in Fig. 11(b) and 11(c). The coding gains of the individual levels are

$$G_1 = \frac{2 + c + d}{4\sqrt{\frac{c+d}{2}}}, \quad G_2 = \frac{c + d}{2\sqrt{cd}}$$

The total coding gain of the tree, given by  $G_1 G_2^{1/2}$ , is therefore equal to (11) showing that the tree achieves the upper bound on the coding gain. This means in particular that the choice of filters shown in Fig. 11 results in an optimal orthonormal two-level dyadic tree. Next, Fig. 11(d) shows the effective filter  $H_{10}(z)H_{20}(z^2)$  of the top channel (which has effective decimator  $\downarrow 4$ ) of the three channel filter bank. Since  $c < 1$ , this is clearly not an optimum compaction(4) filter for the input psd. This example shows that, even though the coding gain is optimized, the top filter  $H_{10}(z)H_{20}(z^2)$  is not an optimum compaction filter for the input  $x(n)$ .

### Nonoptimality of Individual Sections

We pointed out that if a tree structured filter bank is optimal, then the right-flushed subtrees (Fig. 10) are optimal for their inputs. However, the left-flushed subtrees are not necessarily optimal. For example, the gain  $G_1$  of level 1 may not be optimal for the primary input  $x(n)$ . In fact, if we optimize each stage for its input by proceeding from left to right, it may not yield an optimum tree. This is demonstrated with the example shown in Fig. 12, for a two-level dyadic tree. The choice of filters shown earlier in Fig. 11 yields the coding gain (11), and therefore represents an optimal orthonormal filter bank. Instead of using this filter bank, suppose we first optimize the coding gain  $G_1$  of level 1. The unique ideal level-1 filter bank for this is shown in Fig. 12(b), and has coding gain  $G_1 = (2 + c + d)/2\sqrt{(1+c)(1+d)}$ . The decimated subband psd at level 1 are shown in Fig. 12(c), and 12(b). The optimum level 2 filter bank yields the coding gain  $G_2 = (1 + d)/2\sqrt{d}$  (assuming this is greater than  $(1 + c)/2\sqrt{c}$ ) so that the overall gain is

$$G_1 G_2^{1/2} = \frac{2 + c + d}{4(cd)^{1/4}} \times \underbrace{\left( \frac{2\sqrt{c}}{1 + c} \right)^{1/2}}_{\alpha}$$

which is the optimal gain (11) reduced by the factor  $\alpha$ . Since  $\alpha < 1$  unless  $c = 1$ , this filter bank is not optimal.

### Principal Component Property And Compaction

As shown earlier, unlike in uniform filter banks, the coding gain and compaction gain are not directly related in the case of nonuniform orthonormal filter banks. However, there is a simple direct relation between optimal compaction and the principal component property even in the nonuniform case.

In our discussion we find it convenient to change the normalization convention for the filters. Thus consider the example of a 3-level tree of the form Fig. 9. Assuming that the filter bank is orthonormal we have  $\sigma_{x_0}^2 + \sigma_{x_1}^2 + 2\sigma_{x_2}^2 + 4\sigma_{x_3}^2 = 8\sigma_x^2$ . We would like to rescale the analysis filters such that under the new scaling convention  $\sigma_{x_0}^2 + \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_3}^2 = \sigma_x^2$ , or more generally,  $\sum_{k=0}^{M-1} \sigma_{x_k}^2 = \sigma_x^2$ . This is accomplished simply by dividing each two-channel analysis filter  $H_{km}(z)$  by  $\sqrt{2}$  and multiplying the corresponding synthesis filter by  $\sqrt{2}$ . Under the new scaling convention, the signals  $y_k$  indicated in Fig. 9 have variances given by

$$\begin{aligned}\sigma_{y_1}^2 &= \sigma_{x_0}^2 + \sigma_{x_1}^2 \\ \sigma_{y_2}^2 &= \sigma_{x_0}^2 + \sigma_{x_1}^2 + \sigma_{x_2}^2 \\ \sigma_{y_3}^2 &= \sigma_{x_0}^2 + \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_3}^2 \\ &\vdots\end{aligned}$$

If the tree structured filter bank has the principal component property for the given input  $x(n)$ , then the partial sum

$$\sigma_{y_L}^2 = \sum_{k=0}^L \sigma_{x_k}^2$$

is maximized for each value of  $L$  in  $0 \leq L \leq M-1$ . That is, the left-flushed subtrees indicated in Fig. 13 should be such that their top filters

$$H_{10}(z)H_{20}(z^2) \dots H_{i+1,0}(z^{2^i})$$

are optimal compaction( $2^{i+1}$ ) filters for the primary input  $x(n)$ . For example,  $H_{10}(z)$  should be the optimum compaction(2) filter for  $x(n)$ , whereas  $H_{10}(z)H_{20}(z^2)$  must be the optimum compaction(4) filter for the same  $x(n)$  and so forth.

To explain how the above principal component property can be satisfied, consider the example of a three-level tree of the form shown in Fig. 9. First design an optimum compaction(8) filter for the input  $x(n)$ . As shown in<sup>18</sup>, this filter can always be implemented in the multirate-cascade form shown in Fig. 14 where each  $H_{k0}(z)$  is an optimal compaction(2) filter for its input. Since each  $H_{k0}(z)$  is such that  $|H_{k0}(e^{j\omega})|^2$  is Nyquist(2), we can always define a filter  $H_{k1}(z)$  such that the pair  $\{H_{k0}(z), H_{k1}(z)\}$  is a two-channel orthonormal filter bank. Since each  $H_{k0}(z)$  is an optimum compaction filter for its input, the pair  $\{H_{k0}(z), H_{k1}(z)\}$  maximizes the coding gain for its input. In this way, the complete tree structure is defined (by Fig. 9) and satisfies the principal component property. The coding gain of this principal component filter bank, however, is not necessarily maximized as explained earlier.

### 5. CONCLUDING REMARKS

We now summarize a number of related results derived in<sup>17</sup>. The compaction gain  $\sigma_y^2/\sigma_x^2$  in Fig. 6 depends on the input psd as well as the filter  $H(e^{j\omega})$ . The maximum compaction gain  $G_{max}(M)$  is such that  $1 \leq G_{max}(M) \leq M$ . We

have  $G_{max}(M) = 1$  if and only if  $S_{xx}(e^{j\omega}) = S(e^{j\omega M})$ , and  $G_{max}(M) = M$  if and only if  $S_{xx}(e^{j\omega})$  has an aliasfree( $M$ ) support. For arbitrary  $S_{xx}(e^{j\omega})$ , the gain  $G_{max}(M)$  is not monotone increasing in  $M$ , though this can be shown to be true for monotone  $S_{xx}(e^{j\omega})$ . Finally, the optimized coding gain  $G_{SBC}(M)$  of the  $M$  band uniform orthonormal subband coder is not necessarily a monotone increasing function of  $M$  even if  $S_{xx}(e^{j\omega})$  is monotone. This is unlike the behavior of the optimal transform coder (KLT).

For fixed number of subbands  $M$ , biorthogonal filter banks can in general provide better coding gain than orthonormal filter banks. In fact, the coding gain of an optimal orthonormal filter bank can almost always be improved by using a half-whitening prefilter  $H(e^{j\omega})$  and a post filter  $1/H(e^{j\omega})$  around that filter bank. The optimal combination of the prefilter  $H(e^{j\omega})$  and the orthonormal filter bank, which maximizes coding gain, is given in<sup>15</sup>. A related question of interest is, what is the class of input psd  $S_{xx}(e^{j\omega})$  for which an optimal orthonormal filter bank is as good as a biorthogonal one (in the coding gain sense)? For uniform filter banks with fixed  $M$  it is shown in<sup>17</sup> that this is the case if and only if the frequency axis  $[0, 2\pi)$  can be partitioned into  $M$  aliasfree( $M$ ) regions such that  $S_{xx}(e^{j\omega})$  is a constant in each region.

### ACKNOWLEDGEMENTS

This work was supported in parts by Office of Naval Research grant N00014-93-1-0231, and Tektronix, Inc.

### REFERENCES

- [1] Akansu, A.N., and Haddad, R.A., *Multiresolution signal decomposition: transforms, subbands, and wavelets*, Academic Press, Inc., 1992.
- [2] P. P. Vaidyanathan, *Multirate systems and filter banks*, Prentice Hall, Inc., 1993.
- [3] H. S. Malvar, *Signal processing with lapped transforms*, Artech house, Norwood, MA, 1992.
- [4] Vetterli, M., and Kovačević, J. *Wavelets and subband coding*, Prentice Hall, Inc., 1995.
- [5] Strang, G., and Nguyen, T. *Wavelets and filter banks*, Wellesley-Cambridge Press, 1996.
- [6] Akansu, A. N., and Liu, Y. "On signal decomposition techniques," *Optical engr.*, vol. 30, July 1991.
- [7] Haddad, R. A., and Park. K. "Modeling, analysis, and optimum design of quantized  $M$ -band filter banks," *IEEE Trans. on Signal Processing*, pp. 2540–2549, Nov 1995 .
- [8] Unser, M. "On the optimality of ideal filters for pyramid and wavelet signal approximation," *IEEE Trans. on Signal Processing*, v. 41., pp. 3591–3596, Dec. 1993 .
- [9] Unser, M. "An extension of the KLT for wavelets and perfect reconstruction filter banks," *Proc. SPIE No. 2034, Wavelet Appl. in Signal and Image Proc.*, pp. 45-56, San Diego, CA, 1993.
- [10] Tsatsanis, M. K., and Giannakis, G. B., "Principal component filter banks for optimal multiresolution analysis," *IEEE Trans. on Signal Proc.*, vol. 43, pp. 1766–1777, Aug. 1995.
- [11] Delsarte, P., Macq, B., and Slock, D. "Signal-adapted multiresolution transform for image coding," *IEEE Trans. on Info. Th.*, vol. 38, pp. 897–904, Mar. 1992.
- [12] Ohno, S., and Sakai, H. "Optimization of filter banks using cyclostationary spectral analysis," *Proc. Int. Conf. ASSP*, pp. 1292–1295, Detroit, May 1995.



- [13] Moulin, P. "A new look at signal-adapted QMF bank design," Proc. ICASSP, pp. 1312-1315, May 1995.
- [14] Xuan, B., and Bamberger, R. H., "MD paraunitary principal component filter banks," preprint.
- [15] I. Djokovic and P. P. Vaidyanathan, "On optimal analysis/synthesis filters for coding gain optimization," IEEE Trans. on Signal Processing, pp. 1276-1279, May 1996.
- [16] P. P. Vaidyanathan, "Theory of optimal orthonormal filter banks," Proc. IEEE Int. Conf. Acoust. Speech, and Signal Proc., Atlanta, GA, May 1996.
- [17] P. P. Vaidyanathan, "Theory of optimal orthonormal subband coders," IEEE Trans. Signal Processing, submitted.
- [18] Y-P Lin and P. P. Vaidyanathan, "Considerations in the design of optimum compaction filters for subband coders," Proc. Eusipco, Trieste, Italy, 1996.
- [19] Jayant, N. S., and Noll, P. *Digital coding of waveforms*, Prentice Hall, Inc., Englewood Cliffs, 1984.
- [20] P. P. Vaidyanathan, "Orthonormal and biorthogonal filter banks as convolvers, and convolutional coding gain," IEEE Trans. on Signal Processing, pp. 2110-2130, June 1993.

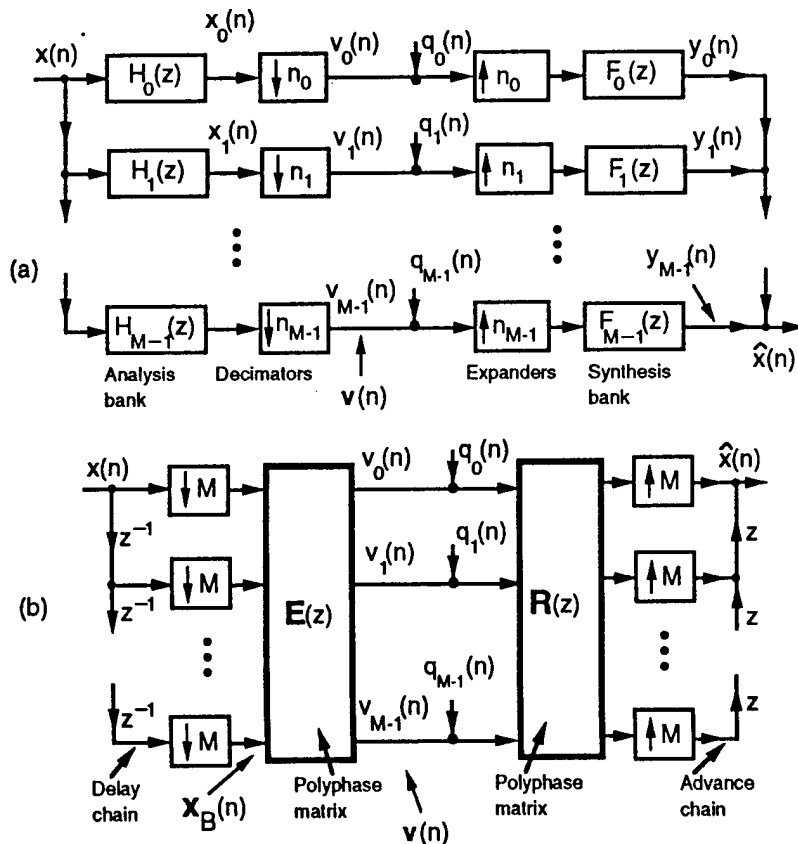


Fig. 1. (a) The maximally decimated subband coder schematic, and (b) the polyphase representation in the uniform case.

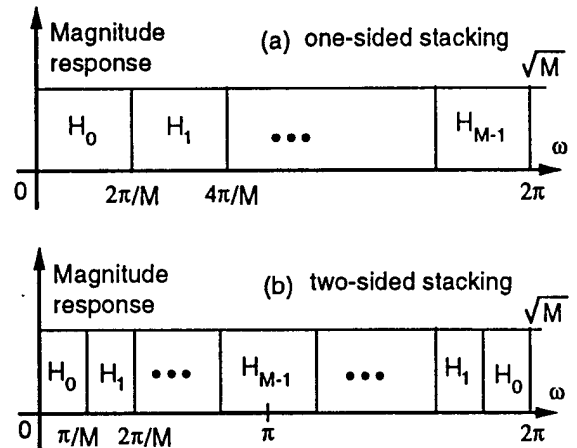


Fig. 2. The brickwall filter bank with contiguous stacking. (a) One-sided stacking, and (b) two-sided stacking as in the real-coefficient case.

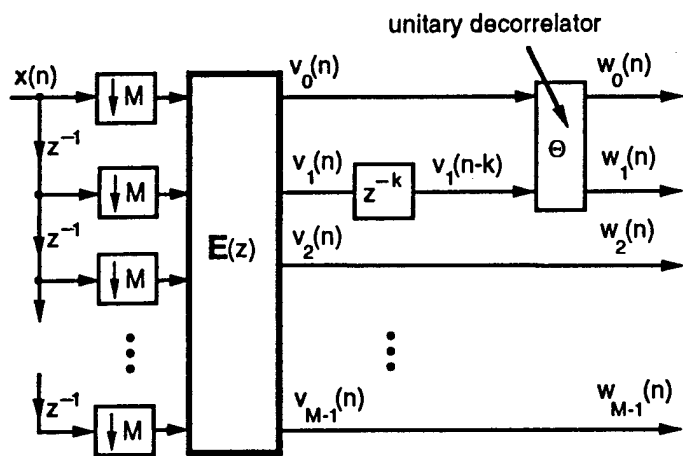


Fig. 3. Increasing coding gain by exploiting correlation.

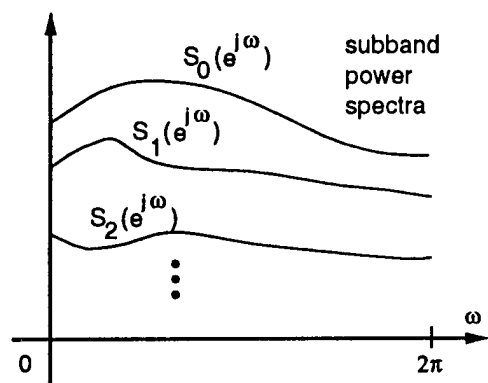


Fig. 4. Demonstration of the majorization property of the decimated subband signals.

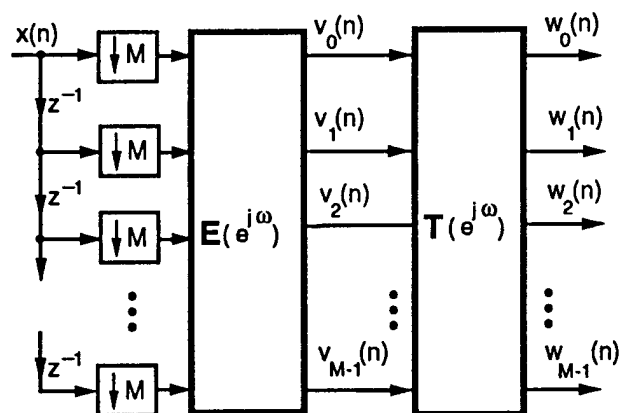


Fig. 5. Increasing coding gain by exploiting unmajorized subband psd.

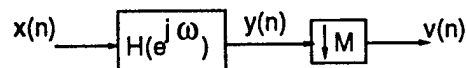


Fig. 6. The energy compaction filter.

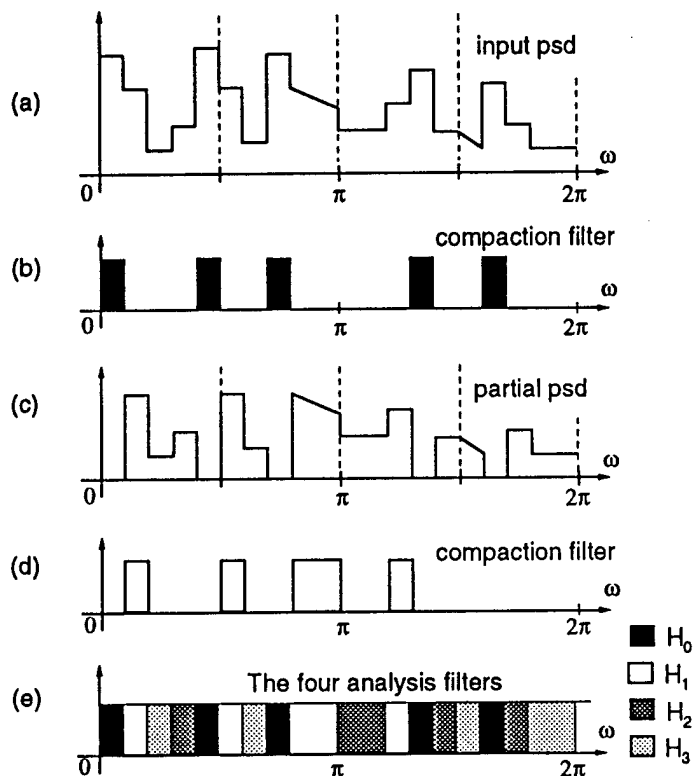


Fig. 7. An input psd and the construction of a 4-band optimal orthonormal filter bank.

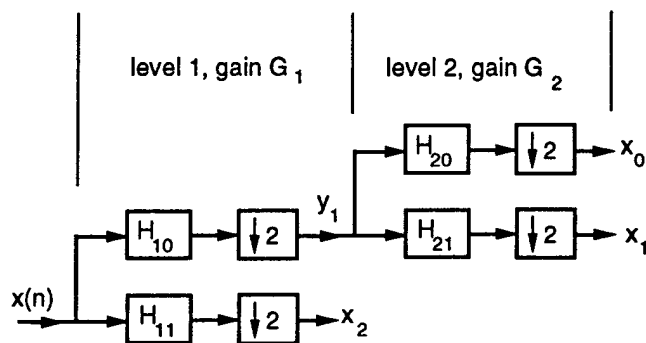


Fig. 8. A two-level dyadic tree structured filter bank.

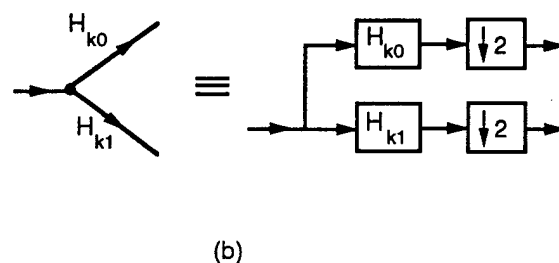
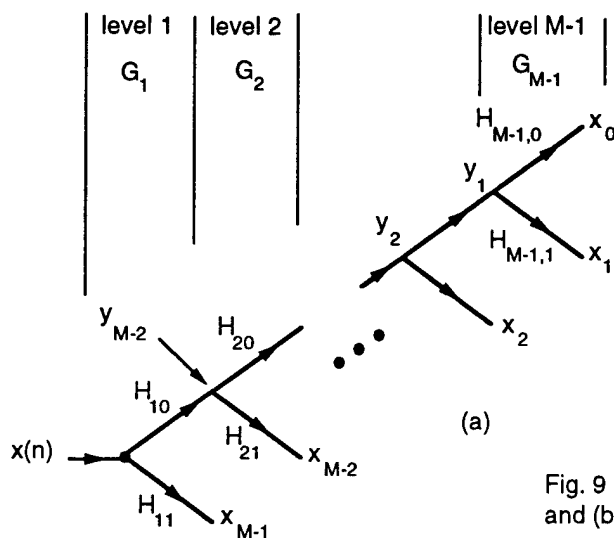


Fig. 9 (a) The general dyadic tree structure schematic, and (b) details of the notation.

Examples of right-flushed subtrees

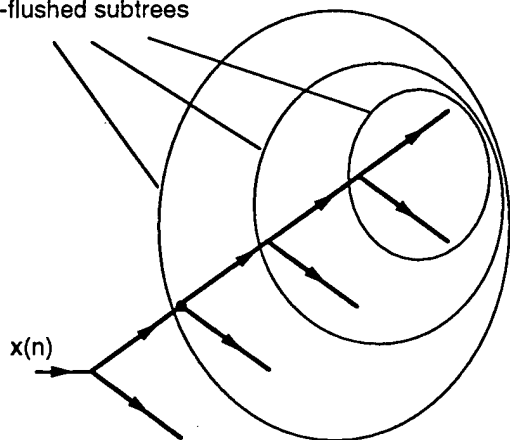


Fig. 10. The right-flushed subtrees of an optimal orthonormal FB are optimal for their inputs.

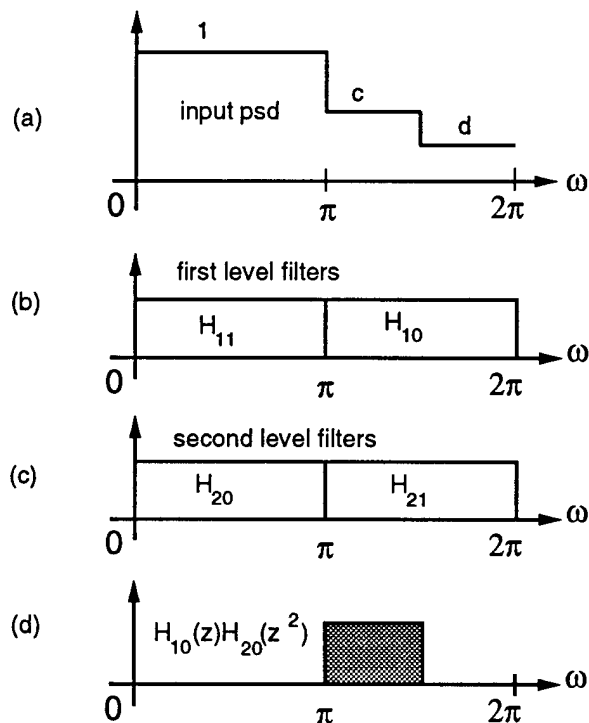


Fig. 11. An input psd and the various filters in a two level tree structured optimal orthonormal filter bank

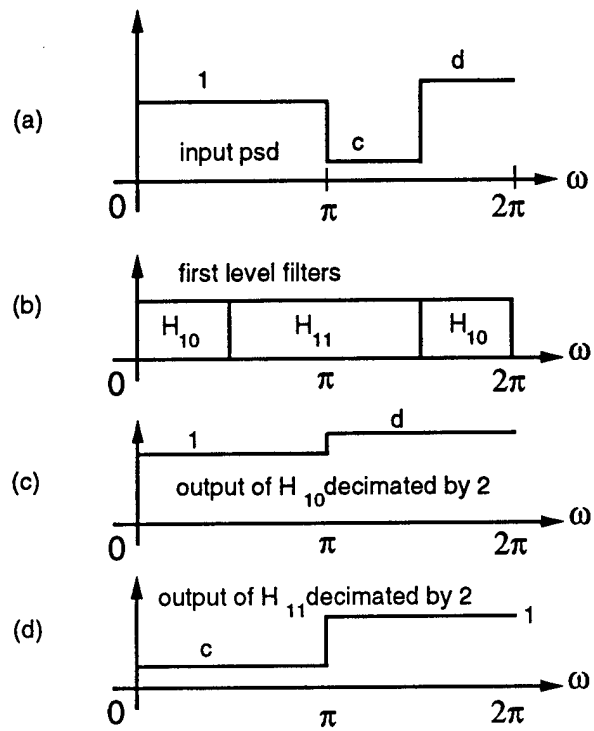


Fig. 12 (a) An input psd, (b) the optimal level 1 filter, and (c), (d) the decimated subband power spectra at level 1.

Examples of  
left-flushed subtrees

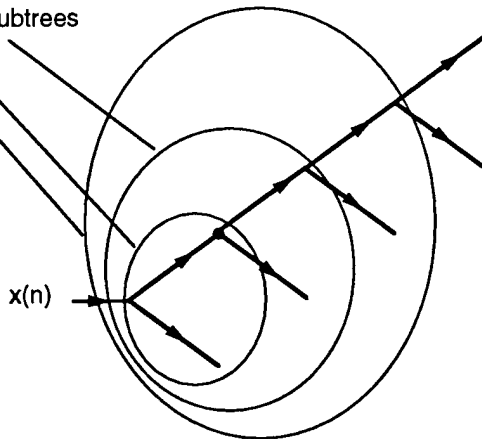


Fig. 13. The left-flushed subtrees of a principal component FB define several optimal compaction filters for the primary input  $x(n)$ .

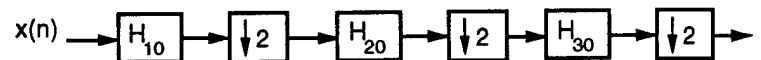


Fig. 14. The cascade implementation of an optimal compaction(8) filter. This completely defines the 3-level dyadic principal component filter bank.